DOCUMENT RESUME

ED 301 475 SE 050 227

AUTHOR Brown, William R.
TITLE Test Mapping.
PUB DATE 88
NOTE 20p.; May not reproduce well. Best copy available.
PUB TYPE Reports - Descriptive (141) -- Guides - Classroom Use - Guides (For Teachers) (052) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Evaluation Methods; Higher Education; *Student Evaluation; Teacher Made Tests; *Test Construction; *Test Format; *Testing; Test Reliability; Test Validity

ABSTRACT

The evaluation tools written by teachers are rarely valid or reliable. One teaching aid that can help in the creation of an effective evaluation instrument is called a test map. A test map is a systematic method to consider variables that are important in the construction of the format of a test. Five variables that are discussed in the test mapping discussions in this paper are time on task, level of learning; type of questions; number of questions; and difficulty order by area. Also included in this document are forms that can be used in this method and an example of the development of a test map using arithmetic content. (CW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED301475

TEST MAPPING

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William R. Brown

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

William R. Brown, Ph.D.

Science Education Center
School of Education
Old Dominion University
Norfolk Virginia 23508

Copyright © 1988 by William R. Brown. All rights reserved. No part of this document may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission of the copyright holder.

SE 050 227

BEST COPY AVAILABLE

"Oh no! It's 10:05 p.m. and I need to write a test for my classes to be given tomorrow! I hope I get to see the 11:00 news. No problem, I'll just sit down and make my test. Let's see, multiple choice are nice. How about ten questions at four points each. Matching are easy to write. Fifteen at two points each will do. Maybe a few true or false will work. Ten at two points each will give me a total of ninety points. I need to finish in 50 minutes! A ten point essay will finish my test at 100 points.

I'll use a few questions from old tests and write a few more. I will write directly on the ditto master and run off copies tomorrow morning. Great, I am all done and it is only 10:55!" Does this scenario sound familiar?

This "measuring device" will be used to measure student learnings. The student's score will be added to other scores to arrive at a final grade for the year. The discriminatory symbol, or grade, will be a matter of record that will follow a person forever! Society accepts these symbols as being equivalent and as a measure of achievement. How can tests generated in the manner described above yield valid and reliable scores? The answer to the question is that evaluation tools written by teachers rarely are valid and/or reliable!

Three steps aid in the evolution of valid and reliable tests. Test mapping, the first step, is discussed in this document. The writing of test questions in a consistent

pattern and a procedure for test analysis are described elsewhere (Brown, 1985).

A test map is a systematic method to consider variables that are important in the construction of the format of a test. Five variables will be used in this example. These are: (1) time-on-task, (2) level of learning, (3) type of questions, (4) number of questions, and (5) difficulty order by area. The map will be constructed using the chart shown in Figure 1.

Time-on-Task

When a teacher spends a lot of time on a topic, the students generally assume that this topic is important. A small amount of time sends a message of lesser importance. To calculate time-on-task, the teacher needs to keep time records and/or make estimates. For example, a two week test could include 700 minutes of learning time (10 days X 50 minutes = 500 minutes in school plus 10 days X 20 minutes = 200 minutes out-of-class). Assume four major areas were covered in this 700 minutes. The teacher calculates that 200 minutes were spent on topic A, 100-topic B, 50-C, and 350-D. This converts to $\frac{200}{700}$=29%, $\frac{100}{700}$=14%, $\frac{50}{700}$=7%, and $\frac{350}{700}$=50%. Topics A-D in English might be: (A) Wordsworth, (B) Coleridge, (C) Keats, and (D) Byron and Shelly. An example from mathematics might be: (A) buying a car, (B) car insurance, (C) depreciation, and (D) repair and

Figure 1: Test Map

	Areas (Time-on-Task)			
Domain ↓	A = ____%	B = ____%	C = ____%	D = ____%
Knowledge				
____% { Comprehension				
Application				
Analysis				
____% { Synthesis				
Evaluation				
____ questions				
__/__ Domains				

upkeep. The percentages calculated for time-on-task are entered on the test map. See Figure 2.

Level of Learning

If we teach "define" and test "compose," the teaching-evaluating process is inconsistent. If we specify "compose" in our objectives and teach "define," a second inconsistency develops. If we specify "compose" in our objectives, teach "compose," and test "compose," we are consistent!

"Compose" is an example from the synthesis level of Bloom's Taxonomy (Bloom, 1956). "Define" is from the knowledge level. Let's assume that 20% of learning time in each area, (A-D), was spent on higher level objectives such as application, analysis, synthesis, and evaluation. Let's make sure that the test reflects this emphasis. The 80%/20% emphasis is entered on the map as shown in Figure 3.

Type of Questions

The teacher wants students to learn to organize ideas. This goal should be reflected in classroom measuring devices. Essay questions, where students must provide the words, are a tool that can serve as an incentive to the student in the area of organizing ideas. However, long essays severely limit the scope of a test. In order to measure facts efficiently, quick response items such as multiple choice and matching are appropriate. In the

Figure 2: Time-on-Task

	Areas (Time-on-Task)			
Domain ↓	A = 29 %	B = 14 %	C = 7 %	D = 50 %
Knowledge				
___% Comprehension				
Application				
Analysis				
___% Synthesis				
Evaluation				
_____ questions				
__/__ Domains				

Figure 3: Level of Learning

	Areas (Time-on-Task)			
Domain ↓	A = 29 %	B = 14 %	C = 7 %	D = 50 %
Knowledge				
80 % { Comprehension				
Application				
Analysis				
20 % { Synthesis				
Evaluation				
_____ questions				
80/20 Domains				

construction of this test map we want to include both short essay (1-10 words) and multiple choice. The type of questions selected affects the number of questions on a test.

Number of Questions

The number of questions on a test should be a function of how long it takes to respond to each item. Teachers can make estimates based on previous tests or they can actually "experiment" with their students and collect data. Let's assume that a student can respond to one well-written multiple choice quesion or three short answer questions in one minute. Let's also specify that the test must be completed in 40 minutes. Thirty-six multiple choice questions would require 36 minutes. Twelve short answer questions would use 4 minutes of test taking time. This yields 48 questions (36 mc and 12 sa) in the 40 minutes. These figures can be added to the test map. See Figure 4.

Difficulty Order by Area

A test should be an evidence gathering device used to see if material related to an objective has been learned. Tests should be a learning tool not a punishment! A way to help a test taker build his self confidence is to place easier items first and save the harder items for the end of the test. Teachers who have obtained item difficulty information by doing a test item analysis can intentionally

Figure 4: Number and Type of Questions

	Areas (Time-on-Task)			
Domain ↓	A = 29 %	B = 14 %	C = 7 %	D = 50 %
Knowledge				
80 % { Comprehension				
Application				
Analysis				
20 % { Synthesis				
Evaluation				
48 questions 36 mc (12 sq)				
80 /20 Domains				

select easy-medium-hard items to be placed appropriately throughout the test.

Another way to place items by difficulty is to estimate which area (A-B-C-D) is easiest for most students. Let's estimate that area B is easiest, area D is second easiest, area C is third easiest, and area A is the hardest. This information is entered on the test map as shown in Figure 5.

The difficulty order is shown on the test map, but it is a teacher decision as to how many easy-medium-hard items will be used. For example, if the teacher wishes to have an "easy" test to help class morale, the test might include 40 easy items and eight of medium level difficulty.

Arithmetic

The basic data for the test map is all shown in Figure 5. It is now time to inter-relate the variables.

First time-on-task by area will be related to the total number of questions. Area A should represent 29% of the 48 questions. Therefore 14 questions on the test will be from this area. Area B: (14%)(48 questions)=7; area C: (7%)(48 questions)=3; and area D: (50%)(48 questions)=24. The number of questions by area are shown in Figure 6.

The second relationship to be calculated is the number of questions per area as related to level of learning. For each area 80% of the questions are "low level" and 20% are from the "higher level" of Bloom's Taxonomy. For area A, 80% of 14 questions are low level and 20% of 14 questions

Figure 5: Difficulty Order by Area (1) is easiest, (4) is hardest.

	Areas (Time-on-Task)			
Domain ↓	A = 29 % (4)	B = 14 % (1)	C = 7 % (3)	D = 50 % (2)
Knowledge				
80% { Comprehension				
Application				
Analysis				
20% { Synthesis				
Evaluation				
48 questions 36 mc (12 sq)				
80/20 Domains				

Figure 6: Number of Questions by Area

	Areas (Time-on-Task)			
Domain ↓	A = 29 % (4)	B = 14 % (1)	C = 7 % (3)	D = 50 % (2)
Knowledge				
80% Comprehension				
Application				
Analysis				
20% Synthesis				
Evaluation				
48 questions 36 mc (12 sq)	14	7	3	24
80/20 Domains				

are higher level. This gives an 11/3 split for lower/higher levels. Area B: (80%)(7 questions)=6/1; Area C: (80%)(3 questions)=2/1; and Area D: (80%)(24 questions)=19/5. Number of questions by domain are shown in the last row on Figure 7.

The Map

Notice on Figure 7 that all the peripheral information is complete. It is now time to place individual test items on the map.

Start with area B since it is the easiest. Test items 1-6 are low level and item 7 is higher level. You write or select items from your test file that are easy and reflect the level of learning. Questions 1, 2, 3, 4, and 6 are mutliple choice. Items 5 and 7 are short answer. Figure 8 shows question placement and type of question for area B. Notice that short answer questions are circled.

The second area by difficulty is area D. Nineteen lower level and five higher level questions are placed on the map to include 18 multiple choice and 6 short answer. Completed areas B and D are shown in Figure 9.

The completed map is shown in Figure 10. Note that twelve questions are circled (short answer) and 36 questions are multiple choice. All the multiple choice questions should be grouped together in the sequence 1, 2, 3, 4, 6, 8-24, 30, 33...47. The short answer questions should be

Figure 7: Number of Questions by Domain

	Areas (Time-on-Task)			
Domain ↓	A = 29 % (4)	B = 14 % (1)	C = 7 % (3)	D = 50 % (2)
Knowledge				
80% Comprehension				
Application				
Analysis				
20% Synthesis				
Evaluation				
48 questions 36 mc (12 sq)	14	7	3	24
80/20 Domains	11/3	6/1	2/1	19/5

Figure 8: Adding Questions to Map

	Areas (Time-on-Task)			
Domain ↓	A = 29 % (4)	B = 14 % (1)	C = 7 % (3)	D = 50 % (2)
Knowledge		1, 2, 3		
80% { Comprehension		4, (5) 6		
Application				
Analysis				
20% { Synthesis		(7)		
Evaluation				
48 questions 36 mc (12 sq)	14	7	3	24
80/20 Domains	11/3	6/1	2/1	19/5

Figure 9: Questions for Second Easiest Area

Domain ↓	Areas (Time-on-Task) A = 29 % (4)	B = 14 % (1)	C = 7 % (3)	D = 50 % (2)
Knowledge		1, 2, 3		8-17
80% Comprehension		4, (5) 6		18-24, (25), (26)
Application				(27), (28)
Analysis				(29)
20% Synthesis		(7)		30
Evaluation				(31)
48 questions 36 mc (12 sq)	14	7	3	24
80/20 Domains	11/3	6/1	2/1	19/5

Figure 10: Completed Map

	Area (Time-on-Task) ④ = easiest; ① = hardest area			
Domain ↓	A = 29 % ④	B = 14 % ①	C = 7 % ③	D = 50 % ②
Knowledge	35-40	1, 2, 3		8-17
80% Comprehension	(41), 42, 43, 44, (45)	4, (5), 6	(32), 33	18-24, (25), (26)
Application	46			(27), (28)
Analysis				(29)
20% Synthesis	47	(7)	34	30
Evaluation	(48)			(31)
48 questions 36 mc (12 sg)	14	7	3	24
80/20 Domains	11/3	6/1	2/1	19/5

-16-

placed together in the sequence 5, 7, 25-29, 31, 32, 41, 45, 48.

This test map shows a logical and consistent development of a test based on five variables. Other variables might be used. The frequency of items by difficulty level such as 1/3 easy, 1/3 medium, and 1/3 hard is an example. A teacher who has established a "test item bank" may want to make sure that at least 10% of the items on a test are "new" and not from the bank. In a sequentially developing subject such as arithmetic it might be desirable to insure that 20% of the questions are from past work. Teachers are encouraged to include their own variables on a test map.

Mapping a test is certainly more time consuming than the 10:05 to 10:55 method described earlier. However, the map, in conjunction with well written items, will produce a more consistent and logical measuring device. A discriminatory symbol (grade) is only meaningful if it is based on a valid and reliable measuring tool!

References

Brown, William R., Handbook for Elementary and Middle School Science Education, 2nd edition, OBSERVE, Inc., P.O. Box 62078, Virginia Beach, VA 23462-0078, ©1985.

Bloom, Benjamin S., et.al., Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I: Cognitive Domain, Longman, Inc., 1956. Reprinted in Brown, 1985 with permission of Longman.